Behavorial and emotional problems in young preschoolers: cross-cultural testing of the validity of Child Behavior Checklist/2-3

Koot, H.M.; van den Oord, E.J.C.G.; Verhulst, F.C.; Boomsma, D.I.

published in
Journal of Abnormal Child Psychology
1997

DOI (link to publisher)
10.1023/A:1025791814893

document version
Publisher's PDF, also known as Version of record

Journal of Abnormal Child Psychology, Vol. 25, No. 3, 1997, pp. 183-196

Behavioral and Emotional Problems in Young Preschoolers: Cross-Cultural Testing of the Validity of the Child Behavior Checklist/2-3

Hans M. Koot,[1,3] Edwin J. C. G. Van Den Oord,[1,2] Frank C. Verhulst,[1] and Dorret I. Boomsma[2]

Received February 14, 1996; revision received July 3, 1996; accepted July 3, 1996

The cross-cultural validity of the Child Behavior Checklist for Ages 2-3 (CBCL/2-3) was tested in three Dutch samples of children referred to mental health services, from the general population, and from a twin study. Six scales were derived from factor analyses and labeled Oppositional, Aggressive, and Overactive, which constituted a broadband Externalizing grouping; Withdrawn/Depressed and Anxious, which constituted a broadband Internalizing grouping; and Sleep Problems. Internal consistencies of the scales, their test-retest reliabilities, interparent agreement, discriminative power, predictive relations with problem ratings 2 years later, and relations to other instruments designed to measure general development and behavior problems were adequate, and highly comparable to psychometric properties in American samples. It was concluded that across languages and cultures behavioral/emotional problems of young preschoolers may be adequately assessed with the CBCL/2-3.

KEY WORDS: Behavior problems; emotional problems; CBCL/2-3; preschool children.

In the last two decades the interest in the psychosocial development of young children has increased considerably. Not only have new theories and fields of research on the social and emotional development emerged (e.g., attachment theory and research; theories of self-regulation), also behavioral and emotional problems specific to this age range have received more attention than before (e.g., Campbell, 1990, 1995; Richman & Lansdown, 1988; Trad, 1989). However, in the study of psychopathology in very young children, basic issues of assessment, taxonomy, and epidemiology of deviant behavior and development have only been scantily addressed. While for older age groups numerous efforts have been undertaken to develop standardized assessment procedures and taxonomic systems of psychopathology (e.g., Achenbach, 1993; Rutter, Tuma, & Lann, 1988), and a considerable number of prevalence studies have been performed involving nearly 250,000 children (see Verhulst, 1995, for an overview), firm data on the prevalence and correlates of symptoms, and on the validity of differentiated syndromes of psychopathology in children below age 4 are still limited.

One reason for the lack of information on psychopathology in young preschool children may be that it is difficult to differentiate between behaviors that can be regarded normal for this age and behaviors that can be considered deviant or pathological. Many behaviors considered problematic in older children, like defiance, anxiety in new situations, or lack of emotional and behavioral control, may be part of normal development in younger children (e.g., Wenar, 1994).

[1]Erasmus University, Rotterdam, The Netherlands.
[2]Free University, Amsterdam, The Netherlands.
[3]Address all correspondence to Hans M. Koot, Sophia Children's Hospital, Department of Child and Adolescent Psychiatry, Dr. Molewaterplein 60, 3015 GJ Rotterdam, The Netherlands.

0091-0627/97/0600-0183$12.50/0

Further, different informants seeing the preschool child under different conditions and having different relations to the child often disagree on the presence and severity of the child's problem behaviors (Spiker, Kraemer, Constantine, & Bryant, 1992). Each informant may provide valid, though different, data. Not surprisingly, no coherent taxonomy exists that might be used to describe behavioral and emotional problems in preschool children. For example, diagnostic categories in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) (DSM-IV; American Psychiatric Association, 1994), such as oppositional disorder or separation anxiety disorder, may have little validity in young preschool children, because many otherwise normally functioning children will qualify for these diagnoses.

Given the problems in defining psychopathology in young children, it may not be surprising that studies on the occurrence of psychopathology in preschool children have yielded widely varying results. The reported prevalence of significant disorder among preschoolers based on parental reports ranges from 6% to 50% (see Koot & Verhulst, 1991, for an overview). This variation in results is likely to reflect methodological differences between studies since prevalence rates were derived from diverse and sometimes unstandardized assessment procedures. Furthermore, a number of studies used unrepresentative samples and many studies differed in their criteria for deviance, i.e., simple percentages of reported behavior, and cutoff points based on either percentile scores or on ratings made by psychiatrists.

Until now, information on the prevalence of problem behavior in young preschoolers was absent in The Netherlands, and a valid instrument to provide such information was not available. Since 1986 a downward extension of the Child Behavior Checklist became available to be completed by parents and caregivers of children aged 2 to 3 years, the Child Behavior Checklist for Ages 2-3 (CBCL/2-3), providing information of six syndromes of psychopathology in young children. Research in the United States (Achenbach, 1992; Achenbach, Edelbrock, & Howell, 1987) indicated good reliability and discriminative validity of the instrument. The instrument was soon translated into Dutch and used in clinical settings. However, normative and psychometric data for the instrument were still absent.

The present study was designed to test the cross-cultural validity of the CBCL/2-3. Replication of the psychometric properties of the American version of the CBCL/2-3 in Dutch samples would support the usefulness of the instrument across different languages and cultures. To this end the Dutch translation of the CBCL/2-3 was applied in three different samples. We assessed the internal consistency of the Dutch problem scales, their test-retest reliabilities, interparent agreement, discriminative power, predictive relation with problem ratings 2 years later, and relations to other instruments designed to measure similar and dissimilar phenomena.

The results of this study were compared with those obtained in American samples. Given the findings from previous American-Dutch cross-cultural comparisons of the CBCL for ages 4 to 16 (Achenbach, Verhulst, Baron, & Althaus, 1987; Verhulst, Achenbach, Althaus, & Akkerhuis, 1988; Verhulst, Akkerhuis, & Althaus, 1985), we expected considerable cross-cultural replicability of the psychometric properties of the CBCL/2-3. However, since principal-components analyses of the instrument's items yielded somewhat different scales in two different American samples (Achenbach, 1992; Achenbach, Edelbrock, & Howell, 1987), we might similarly expect somewhat different composition of the Dutch scales.

METHOD

Measures

CBCL/2-3. The CBCL/2-3 (Achenbach, 1992; Achenbach, Edelbrock, & Howell, 1987) is a 99-item instrument to obtain ratings of behavioral/emotional problems by parents or caretakers of children aged 2 and 3 years. Fifty-nine of the items have counterparts on the Child Behavior Checklist for Ages 4-18 (CBCL/4-18; Achenbach, 1991), while the remaining items have been developed specifically for ages 2 to 3. The CBCL/2-3 requires fifth-grade reading skills to complete. Most respondents can complete the form in less than 10 min. Respondents are requested to rate the items that describe the child now or within the past 2 months as 2 if the item is *very true or often true* of the child, as 1 if the item is *somewhat or sometimes true* of the child, and as 0 if the item is *not true* of the child. On 12 open-ended items the respondent is asked to describe the behavior, making it possible to correct the scoring according to the scoring instructions when necessary, and to prevent more than one item from being scored for the same problem.

CBCL/4-18. The CBCL/4-18 (Achenbach, 1991) has 118 problem items using the same format as the CBCL/2-3. The instrument's reliability and validity (Achenbach, 1991) were confirmed for the Dutch situation (Verhulst et al., 1985). The instrument can be scored on eight small-band syndromes and on Internalizing, Externalizing, and Total Problems scales. Only the latter three scores were used in this study.

Behaviour Checklist. The Behaviour Checklist (BCL; Richman, 1977) for 3-year-olds is modeled after the Behaviour Screening Questionnaire (Richman & Graham, 1971). It consists of 19 items regarding 12 problem areas: eating, soiling, sleeping, activity, concentration, dependency/attention seeking, control, tempers, mood, worries, fears, and behavior with siblings and peers. The parent is asked to choose which behavioral description out of three or four choices best fits their child over the past 4 weeks. The checklist can be completed independently by the parent within 5 min. A total BCL score was computed by summing the highest ratings (0, 1, or 2) in each problem area covered by one or more items.

The BCL has been used as a screening instrument to detect preschoolers at risk for behavior problems (Richman, Stevenson, & Graham, 1982). Richman (1977) reported a 4-week test-retest reliability of $r = .81$. Using a score of 10 points or more, the sensitivity of the BCL was 69.6 and the specificity 87.4 with a clinical rating of mild, moderate, or severe behavior problems as criterion (21.5% misclassifications), which compares to a sensitivity of 90.2 and a specificity of 93.2 for the BSQ (8.3% misclassifications). The BCL was used as a measure for the convergent construct validity of the CBCL/2-3.

Minnesota Child Development Inventory. The Minnesota Child Development Inventory for 0- to 6-year-olds (MCDI; Ireton & Thwing, 1974) asks a true/false response of parents on 320 empirically derived items describing children's development and behavior. The caregiver is asked to respond *yes* or *no* to all of the items, which takes approximately 45 min. The items form eight developmental scales: Gross Motor, Fine Motor, Expressive Language, Conceptual Comprehension, Situation Comprehension, Self-Help, and Personal-Social scales, and a General Development scale, which is composed of the 131 items from the seven other scales best discriminating between age groups. The MCDI has been shown to correlate significantly with objective developmental and language measures. In this study, the MCDI was used as a measure for the divergent construct validity of the CBCL/2-3. The MCDI was translated into Dutch. Because no independent Dutch normative data are available for the MCDI, we constructed sex- and age-related norms. For each sex eight 3-month age groups were formed using the data from the general population sample in this study, and for each group frequency distributions of the scale scores were computed. Using these distributions, the raw scores for each child were substituted by sex- and age-appropriate percentile scores. These percentile scores were used in the analyses.

Subjects and Procedures

The CBCL/2-3 was used in three different samples. Our purpose was to compose CBCL/2-3 scales based on robust factors, i.e., based on factor solutions that are reasonably invariant across variations in the selection of subjects. To the extent that invariance can be found across different samples, the factors have a wider range of applicability as generalized constructs.

Clinical Sample. The clinical sample consisted of 426 children (284 boys, 142 girls) consecutively referred to 12 child guidance and mental health settings for behavioral and emotional problems and developmental delays, in which all types of nonresidential mental health institutions available for 2- to 3-year-olds in The Netherlands were represented. The mean age of the children was 36.1 months ($SD = 8.1$). Ethnicity was 79.9% Caucasian, 8.2% Surinam/Antillean, 3.1% Mediterranean, and 8.8% from other ethnic groups. Parental educational levels were coded according to a 9-step scale (Van Westerlaak, Kropman, & Collaris, 1975), which was recoded for purpose of analysis to a standard Dutch 4-step scale (Central Bureau of Statistics, 1987), where 1 = *elementary school*; 2 = *lower level secundary or professional education;* 3 = *medium level secundary or professional education;* 4 = *university or high level professional education.* The mean educational level of mothers was 2.35 ($SD = 0.84$), and of fathers was 2.55 ($SD = 0.96$). The employment rates were 89% for fathers and 20% for mothers. The occupational levels of parents who were employed were scored on a standard Dutch 6-step scale (Van Westerlaak et al., 1975), where 1 = *unskilled employees*; 2 = *skilled manual employees*; 3 = *clerical, technicians, minor professionals*; 4 = *owners of small businesses*; 5 = *supervisory, lesser*

professionals; 6 = *executives, major professionals, owners of large businesses.* The mean occupational level of mothers was 2.95 (*SD* = 1.44), and of fathers was 3.00 (*SD* = 1.60). The mean maternal age was 30.3 years (*SD* = 5.0), and the mean paternal age was 33.9 years (*SD* = 5.9).

For the clinical sample, the participating services were asked to have parents, or others in custody who came with each child, fill out the CBCL/2-3 as part of the intake procedure. In preparation of the data collection, mental health workers and office personnel who were in some way involved were instructed on the purpose and procedures of the study and on how to help parents complete the checklists. Letters of introduction, including a description of the study, informed consent forms, and CBCL/2-3s, were handed over to the parents and caretakers at intake. CBCL/2-3s were filled out at the office or at home, and checked by the mental health worker. In the clinical sample, 66.8% of the respondents were mothers, 6.2% were fathers, and 27.4% were both parents or others who were in custody of the child. Those who completed the CBCL/2-3 but were not biological parents were 19 stepparents, five adoptive parents, three grandparents, three foster parents, and one caregiver in a residential setting. The distribution of CBCL/2-3 scale scores was not different for different informants. Demographic information was obtained from the clinical files of the children.

Community Sample. To obtain a representative sample of Dutch 2- to 3-year-olds, names and addresses were drawn from the inoculation register of the province of Zuid-Holland, which includes 95% of all children aged 2-3 years living in the province, except those living in Rotterdam. A sample of 400 children (199 boys, 201 girls) was drawn randomly and stratified by age and sex. In addition and proportionally to the rest of the province, 69 children (34 boys, 35 girls) were drawn, stratified by age and sex, from the Rotterdam municipal health service register, which includes all 2- to 3-year-olds living in the municipality. The resulting target sample included 469 children (233 boys, 236 girls) from 109 sites spread over the province. Birth dates were evenly distributed over the months between October 15, 1985, and October 14, 1987.

Data were collected for 420 children (215 boys, 205 girls) from the target sample. Response rate corrected for untraceables was 91.1%. The mean age of the children was 36.4 months (*SD* = 7.0). Ninety-five percent of the children were of Dutch origin, 1.6% were of Surinam origin, 0.2% came from the Dutch Antilles, 0.2% were Turkish, and 3.1% had another nationality. The mean educational level (CBS, 1987; see clinical sample) of mothers was 2.56 (*SD* = 0.80), and of fathers 2.74 (*SD* = 0.87). The employment rates were 93% for fathers and 32% for mothers. The mean occupational level (Van Westerlaak et al., 1975; see clinical sample) of mothers was 3.53 (*SD* = 1.41) and of fathers 3.68 (*SD* = 1.44). Mean maternal age was 31.5 years (*SD* = 4.4) and the mean paternal age was 34.2 years (*SD* = 5.0). Ten children (2.4%) had been referred to a child mental health agency within the past 12 months.

In the community sample, a letter was sent to the parents of the 469 eligible children explaining the purpose of the study, the way in which the children were selected, and an announcement that an interviewer would contact them. The parents were contacted by telephone, and subsequently visited by one of four trained female interviewers, who had an education at the master's level in special education or psychology. The interviewer read the CBCL/2-3 problem items regarding the target children aloud, and scored the parents' responses. In all cases the mother was the prime respondent. After completing the CBCL/2-3, the parents were asked questions about demographic characteristics of their families. The duration of each interview was 30-60 min.

Parents of the 420 children who participated in the study at Time 1 were contacted after a mean interval of 2.27 years. Parents were visited by one of four trained female interviewers, blind to the Time 1 data. Usable parent information (98% mothers; 95.7% Caucasian) was obtained for 397 children (204 boys, 193 girls; mean age 5.31 years), i.e., 95.2% of the Time 1 responders or 86.7% of the original target sample. Responders and nonresponders did not differ on Time 1 problems, sex, age, socioeconomic status (SES), or maternal and paternal education.

Twin Sample. Subjects in the twin sample were 1,306 pairs of 3-year-old twin pairs (1,291 boys, 1,321 girls) from a target sample of 1,892 pairs (73% response rate). The twins' mean age was 42.1 months (*SD* = 4.0). The employment rates were 98% for fathers and 29% for mothers. The mean occupational level (Van Westerlaak et al., 1975; see clinical sample) of mothers was 3.60 (*SD* = 1.37) and of fathers 3.51 (*SD* = 1.40). The mean maternal age was 33.0 years (*SD* = 3.9), and the mean paternal age was 35.6 years (*SD* = 4.6).

In The Netherlands, about 85% of the parents of all newborns are paid a home visit by a commercial organization which promotes certain products. During this home visit parents of twins are asked to participate in the twin register kept by the Department of Psychonomics of the Free University of Amsterdam. Forty percent of all multiple births in the Netherlands are registered. CBCLs for children aged 2-3 were mailed to parents of 3-year-old twins. Nonresponders were sent reminders and contacted by telephone. For 73% of the twin pairs, both parents filled out one CBCL/2-3 for each child. For 20%, only maternal ratings were available. For 8%, only paternal ratings were available. Questions about demographic characteristics were contained in the questionnaire.

Statistical Procedure

We performed principal-factors analyses using the SAS Institute's (1989) statistical package to identify CBCL/2-3 syndromes in each of the three samples. To make maximum use of the available information in the twin sample, the mean of the parental ratings on each item was used. When available, missing values for one parent were substituted by the ratings of the other parent.

Two items were reported for less than 5% in all samples and were therefore excluded from the analyses: Item 39, Headaches, and Item 57, Problems with eyes without medical cause. The remaining 97 items were subjected to principal-factors analysis using unweighted least squares. The first 5 to 12 factors from each analysis were subjected to varimax and subsequently to oblique promax rotations. For each sample we examined the 5- to 12-factor solutions to identify sets of items that consistently grouped together across solutions within that sample. The factors chosen from the exploratory analyses were subjected to confirmatory factor analyses (CFA) using LISREL 7 (Jöreskog & Sörbom, 1989) to evaluate their applicability across the three samples. Items with loadings ≥ .30 on the same syndrome in at least two of the three samples were included in the factor model to be evaluated in the CFA.

Pearson correlations and *t*-tests were computed to assess test-retest reliabilities of the resulting scales, interparent agreement, convergent and divergent validity, predictive relations with problem scores across a 2-year interval, and comparison of Dutch and American scale scores. Further, multiple-regression analyses were used to test the power of the scales to discriminate between scores of matched referred and nonreferred samples. Chi-square tests were used in analyses on categories of high- and low-scoring children.

RESULTS

Factor Analyses

In both the clinical sample and the community sample, the first seven factors found in the seven- through nine-factor solutions had nearly exactly the same items loading ≥ .30 on similar factors in consecutive rotated solutions. These factors replicated quite well in seven factors from the nine-factor solution of the twin sample. These factors were deemed to be the most representative versions of these syndromes in that they included the largest proportions of items found to cooccur on factors that appeared in multiple rotations within each sample. Inspection of the factor intercorrelation matrix for the oblique factors showed that the oblique solution was clearly preferable to the orthogonal solution.

The items included in the seven-factor solutions in the clinical and community sample and those in the comparable factors from the nine-factor solution in the twin sample with loadings ≥ .30 were listed side by side. In this way we could identify items that would constitute the syndromes to be further evaluated. Table I (first, third, and fifth columns) contains the loadings (standardized regression coefficients ≥ .30) of the 69 items on the seven corresponding factors and the eigenvalues of the unrotated factors in the three samples.

For these 69 items a loading was specified for the syndrome on which they loaded ≥ .30 in at least two of the three samples. Loadings on the other syndromes were fixed at zero. Fourteen of the 69 items had one or more cross-loadings of .30 to .40 in one of the samples, which were all estimated too. Thus, a total of 83 factor loadings were estimated. Finally, all correlations between the syndromes were estimated in the model to be evaluated in the CFA. Table I displays the factor loadings obtained from fitting the final model to the item correlations in the three samples (second, fourth, and sixth column). Only loadings with an absolute value ≥ .30 are shown. In general, parameter estimates were quite acceptable in all three samples. The goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), and root-mean-squared residuals (RMR) of the model were .93, .93, and .06 for the clinical, .94, .93, and .05 for the community, and .97, .97, and .04 for the twin samples, respectively. All fit indices indicated that the model offered a satisfactory description of the test structure.

Table I. Factor Loadings for CBCL/2-3 Items Obtained from Exploratory and Confirmatory Factor Analyses in the Clinical, Community, and Twin Samples[a]

			Clinical sample		Community sample		Twin sample	
Factor/items			EFA	CFA	EFA	CFA	EFA	CFA
I. Oppositional (17 items)								
	8.	Can't wait	.57	.40	.33	.45	.68	.63
	13.	Cries much	.46	.57	.32	.51	.43	.59
A	15.	Defiant	.45	.57 [4]	.35	.46	.32 [3]	.37 [3]
A	16.	Demands must be met	.63	.74	.37 [5]	.68	.73	.79
A	29.	Easily frustrated	.47	.55	—	.43	.49	.51
A	30.	Easily jealous	.34 [3]	.55	—	.43	.48	.56
AD	33.	Feelings easily hurt	—	.49	.36	.38	.52	.55
D	36.	Gets into everything	—	.48	.36	.38	.42	.47
A	44.	Angry moods	.69	.68	.53	.56	.75	.67
A	66.	Screams	.57	.49	.31	.38	.42	.41
A	69.	Selfish	.32 [3]	.31	.35	.31	.34	.39
W	81.	Stubborn	.72	.82	.67	.72	.73	.89
A	82.	Moody	.54	.55	.51	.38	.62	.55
	83.	Sulks	.48	.64	.59	.67	.69	.79
A	85.	Temper tantrums	.76	.67	.57	.57	.76	.70
A	88.	Uncooperative	.55	.56	—	.56	.30	.64
AD	96.	Wants attention	.43	.52	.37 [5]	.43 [5]	.42	.46
A	97.	Whining	.57	.59	.53	.54	.54	.62
Eigenvalue/Cronbach's alpha			14.6/.90		11.1/.86		15.5/.91	
II. Withdrawn/Depressed (10 items)								
W	2.	Acts too young	.63	.58	.41	.39	.48	.34
W	23.	Doesn't answer	.50	.42	.32	.51	—	.57
W	26.	No fun	—	.57	.30	.42	.34	.35
AD	43.	Looks unhappy	—	.56	.58	.66	.44	.49
	56.	Clumsy	.30 [5]	.26	.35	.29	.42 [5]	.17
W	67.	Unresponsive	.40	.52	.37	.46	.46	.36
W	70.	Little affection	.38 [3]	.39	.47	.54	.63	.29
W	71.	Little interest	.51	.63	.47	.63	.59	.48
	76.	Speech problem	.55	.13	.36	.33	—	.25
	77.	Stares blankly	.41	.46	.48	.53	—	.39
	80.	Strange behavior	—	.38	.41	.45	.36	.42
W	89.	Underactive	—	.35	.35	.28	.42	.28
AD	90.	Sad	—	.28 [4]	.32	.35	.48	.50
W	98.	Withdrawn	.50	.60 [1]	.54 [5]	.63	—	.42 [4]
Eigenvalue/Cronbach's alpha			2.2/.74		3.7/.73		1.9/.64	
III. Aggressive (9 items)								
D	14.	Cruel to animals	.47	.48	.54	.39	.41	.33
D	17.	Destroys own things	.32	.63	.54	.50	.72	.59
D	18.	Destroys other's	.36	.63	.65	.46	.73	.58
A	20.	Disobedient	[.52:1]	..38 [5]	—	.31	.30 [1]	.40
A	35.	Fights	.64	.67	.44	.57	.43	.65
A	40.	Hits	.54	.72	.47	.61	.51	.66
D	42.	Hurts accidentally	.42	.52	.50	.53	.46	.50
	53.	Attacks people	.58	.73	.41	.58	.45	.61
A	91.	Too loud	[.42:1]	.42	[.35:1]	.27 [1]	[.36:1]	.53
Eigenvalue/Cronbach's alpha			5.4/.85		2.6/.76		3.8/.82	

(Continued)

Table I. *Continued*

IV. Anxious (9 items)								
	3.	Afraid to try new	.57	.52	.54	.54	.56	.51
W	4.	Avoids eye contact	—	.06 [2]	.40	.35	.61	.55
AD	10.	Clings to adults	.52 [5]	.58	.43	.52	.53	.59
	21.	Disturbed by change	.37	.41	.32	.33	.38	.38
	32.	Fears	.47	.44	—	.30	.38	.32
AD	37.	Upset by separation	.41	.53	.42	.42	.45	.61
AD	68.	Self-conscious	—	.48	.44	.43	.49	.48
AD	73.	Shy	.48	.59	.61	.61	.74	.72
AD	87.	Too fearful/anxious	.61	.59	.45	.53	.56	.59
	92.	Upset by new	.62	.67	.53	.63	.72	.73
Eigenvalue/Cronbach's alpha				3.4/.79		2.3/.76		2.0/.83
V. Overactive (5 items)								
D	5.	Can't concentrate	.64	.95	.67	.95 [1]	.66	.95 [1]
	6.	Can't sit still	.44 [1]	.83	.65	.79	.49	.79
	11.	Constantly seeks help	.39	.49	.41	.54	.35	.61
D	59.	Shifts activity	.52 [1]	.76	.61	.67	.59	.85
W	62.	Refuses active games	.44	.46	.52	.47	.34	.37
Eigenvalue/Cronbach's alpha				1.7/.77		1.6/.77		1.3/.78
VI. Sleep Problems (7 items)								
SL	22.	No sleep alone	.59	.61	.59	.45	—	.39
SL	38.	Can't sleep	.64	.71	.62	.57	.62	.57
SL	48.	Nightmares	.47	.53	.51	.51	.61	.59
SL	64.	Resists going to bed	.51	.71	.38	.58	.48	.55
SL	74.	Sleeps little	.54	.54	.43	.53	.48	.44
SL	84.	Talks/cries in sleep	.37	.45	.40	.42	.48	.47
SL	94.	Wakes often	.56	.73	.65	.70	.73	.62
Eigenvalue/Cronbach's alpha				2.0/.81		1.5/.74		1.5/.70
VII. Somatic Problems (3 items)								
SO	1.	Aches	.40	.66	—	.58	.50	.61
SO	12.	Constipated	.40	.42	.39	.15	—	.27
SO	45.	Nausea	.42	.39	.46	.36	.54	.47
SO	52.	Painful stools	.47	.26	.37	.33	.34	.26
SO	78.	Stomachaches	.43	.48	.30	.42	.60	.58
SO	93.	Vomiting	.41	.25	.37	.13	.36	.29
Eigenvalue/Cronbach's alpha				1.4/.50		1.4/.43		1.1/.59

[a] CBCL/2-3 = Child Behavior Checklist for Ages 2-3; EFA = exploratory factor analysis; CFA = confirmatory factor analysis. Loadings are standardized regression coefficients obtained from promax rotations for the EFA, and unweighted least-squares LISREL estimates for the CFA. Items that were not retained in the scales are underlined. Cross-loadings are given followed by the number of the factor [in square brackets] on which the particular cross-loading occurred. Abbreviations preceding the items indicate items that are comprised in the CBCL/2-3 syndrome scales constructed by Achenbach (1992): A = Aggressive; W = Withdrawn; D = Destructive; AD = Anxious/Depressed; SL = Sleep Problems; SO = Somatic Problems. Underlined items were not included in the final scales.

Table II. Three-Week Test-Retest Reliabilities' Interparent Agreement of CBCL/2-3 Scale Scores, and Relations with Behavior Problem and Development Scales[a]

	Test-retest (n = 51)	Interparent (n = 60)	RCL (n = 207)	MCDI (n = 391)
Oppositional	.88 B	.64	.59	(-.05)
Withdrawn/Depressed	.60 B	.37	.28	-.15
Aggressive	.85	.56	.40	(-.03)
Anxious	.83 BG	.44	.34	-.13
Overactive	.84 B	.64	.54	-.14
Sleep problems	.76 B	.70	.38	-.14
Internalizing	.81 B	.48	.36	-.16
Externalizing	.90 B	.66	.63	(-.08)
Total Problems	.87 BG	.66	.65	(-.10)
Mean r	.82	.56	.41	(-.06)

[a] All rs significant at $p < .01$, except those between brackets. CBCL/2-3 = Child Behavior Checklist for Ages 2-3; BCL = Behavior Checklist; MCDI = Minnesota Child Development Inventory. B, G: Time 1 > Time 2 scores for boys (B) and/or girls (G) ($p < .05$).

Based on the items included in these factors, we applied the following labels to the factors: Oppositional, Withdrawn/Depressed, Aggressive, Anxious, Overactive, Sleep Problems, and Somatic Problems. The results were quite similar for the three samples. Pearson correlations computed for the pairwise comparisons of factor loadings between samples as measures of the congruity of the syndromes across samples (cf. Tanaka & Huba, 1984) indicated a high mean congruity of .81 (range .18 to .97) for all syndromes except Withdrawn/Depressed, for which the mean congruity was .18. These findings suggested that with the possible exception of Withdrawn/Depressed, the same factor structure applied to all three samples. Based on the factor analyses, scales were constructed from the items that were common to a factor in two out of three samples and labeled according to the factor names mentioned above. An item was only included in the scale when its loading on the factor exceeded .30 in two of the three samples. Further, the item was not allowed to have loadings on more than one other factor above an absolute value of .30. This was determined by computing Cronbach's alpha for the scale with and without the item. Because the Somatic Problems scale was substantially shortened by this procedure, we decided to delete it from further analyses. Cronbach's alpha is given in Table I as an index of the internal consistency of the scales.

Second-Order Groupings of Syndromes

To determine whether the six syndrome scales formed broad-band groupings comparable to those found by Achenbach (1992), we performed EFAs and CFAs similar to those for the syndromes on the pattern of observed and weighted least-squares correlations between the syndrome scales, respectively. Both analyses clearly indicated a two-factor model. The first factor was defined by high loadings in every sample for Aggressive (mean of the standardized loadings M = .75), Oppositional (M = .74), and Overactive (M = .64), and may be labeled Externalizing. The second factor was defined by high loadings for Anxious (M = .88) and moderate loadings for Withdrawn/Depressed (M = .46). This factor may be labeled Internalizing. The largest cross-loading was for Withdrawn/Depressed on the Externalizing group (M = .27). The mean loading of Sleep Problems never exceeded .29 on either the Externalizing or Internalizing grouping.

Reliability and Interrater Agreement

To calculate test-retest correlations, the CBCL/2-3 was completed twice over a mean interval period of 19.4 days (SD = 6.6) by 51 respondents (49 mothers, 2 fathers) randomly selected from the community sample. As shown in Table II, test-retest rs ranged from r = .94 for Total Problems in girls to r = .59 for the

Table III. Percentage of Variance Accounted for by Significant Effects of Referral Status, Sex, Age, and Socioeconomic Status (SES) in Scale Scores for Matched Referred and Nonreferred Samples[a]

Scale	Referral status	Sex	Age	SES	Odds ratio
Oppositional	19	—	—	3	12.6
Withdrawn/Depressed	23	—	1	< 1	11.2
Aggressive	14	4 B	—	3	6.6
Anxious	7	—	1	1	5.2
Overactive	17	—	—	2	5.1
Sleep problems	9	< 1 G	—	< 1	5.9
Internalizing	17	—	1	1	6.3
Externalizing	22	< 1 B	—	4	7.8
Total Problems	27	—	—	4	8.7

[a] n = 652 (396 boys; 256 girls) matched by sex, age, and SES. All problem scores were higher for referred than nonreferred in significant cases at p = .00001. All significant age effects reflected higher scores for 3-year-olds ($p < .05$). B = higher scores for boys; G = higher scores for girls ($p < .05$). All significant SES effects reflected higher problem scores for lower than upper SES ($p < .05$). All odds ratios indicated that the proportion of referred children scoring above cutpoint was significantly higher than the proportion of nonreferred ($p < .01$).

Withdrawn/Depressed scale in girls (mean r = .82). All rs were significant at $p < .001$ and above .75, except for the Withdrawn/Depressed scale. Of the 10 comparisons between the reliability coefficients of boys and girls, only one yielded a significant difference. Girls' total problems score appeared to be more reliable than boys' (rs .94 vs. .75, $z = 2.56$, $p = .01$).

Because it is not only important to know rank-order stability of the scale scores as expressed by the Pearson correlation, but also to know changes in magnitude of the scores over a short interval, we computed t-tests between the same ratings for boys and girls separately. As indicated in Table II, eight scores declined significantly ($p < .05$). Over the 3-week interval, for boys all scores except Aggressive showed significant declines ($p < .05$). For girls, however, only the anxious scale score and the Total Problems score declined significantly.

To assess interparent agreement, we computed Pearson correlations and t-tests between scores on CBCL/2-3s independently completed by both parents for 48 cases randomly selected from the community sample and clinical sample. All correlations between parents' scores were significant at $p < .05$ (see Table II). For Sleep Problems, the interparent agreement was significantly higher for girls than for boys ($z = 2.48$, $p < .05$). No significant differences in levels of scale scores between mothers and fathers were found.

Discriminative Validity

To test the discriminative validity of the Dutch CBCL/2-3 scales, we compared demographically matched referred and nonreferred samples. We drew 326 subjects from the clinical sample of 458 subjects who could be matched to 326 subjects from the community sample of 410 subjects (excluding the 10 referred children in this sample). The subjects were exactly matched for sex (198 boys, 128 girls in each group), age (144 of age 2, 182 of age 3 in each group), and SES (111 from lower-SES groups, 121 from middle-SES groups, and 94 from higher-SES groups). For the nonreferred sample, 99.1% of CBCL/2-3s were completed by mothers, 0.6% by fathers, and 0.3% by other informants, compared to 62.6%, 6.1%, and 30.4%, respectively, for the referred sample. Multiple regressions were performed for all problem scales and for the Total Problems score for each sex separately. Table III displays the percentage of variance in each scale accounted for by significant associations with referral status, sex, age, and SES. The percentage of variance accounted for by each variable is represented by the semipartial R^2 (or part correlation) obtained from the regression analyses after partialing out the effects of any other independent variables that accounted for more variance in the scale scores.

Table IV. Correlations of CBCL/2-3 Scores with Corresponding CBCL/4-18 Scores[a]

	Boys		Girls	
Scale	Age 2 (n = 94)	Age 3 (n = 97)	Age 2 (n = 97)	Age 3 (n = 109)
Internalizing	.34	.51	.46	.33
Externalizing	.52	.47	.54	.63
Total Problems	.53	.57	.60	.69

[a] CBCL/2-3 = Child Behavior Checklist for Ages 2-3; CBCL/4-18 = Child Behavior Checklist for ages 4-18. All rs significant at $p < .001$.

All problem scales were scored higher for referred than nonreferred children ($p < .001$). These effects ranged from 7% for the Anxious scale to 27% for the Total Problems score. A few significant sex differences appeared, indicating that boys had higher Aggressive and Externalizing scores than girls, and girls had somewhat higher scores on Sleep Problems. Significant age effects indicated a small increase of Withdrawn/Depressed, Anxious, and Internalizing scores with age. Children from lower-SES strata had significantly higher scores on all scales than children from higher-SES strata. Although significant ($p < .05$), all age, sex, and SES effects were small according to Cohen's (1988) criteria.

For practical purposes, such as clinical decision making, some users may wish to use the CBCL/2-3 scores in a categorical fashion, e.g., by distinguishing children scoring above a certain cutoff point from those scoring below it. To test the discriminative validity of categories based on cutpoints, children from the matched referred and nonreferred samples were classified on each scale according to whether they scored in the "normal" range, below the cutoff, or above it. Cutoff points were based on the 85th percentile of the distribution of Internalizing, Externalizing, and Total Problems scores in the community sample, and on the 95th percentile point for the other scales. Then, the odds were computed that children who had scores above the cutoff point on a particular scale were from the referred sample, relative to the odds for children who did not have scores above the cutpoint. The largest odds ratios (OR, see Table III) were for Oppositional (OR = 12.6) and Withdrawn/Depressed (OR = 11.2), indicating that the odds of being in the referred group were more than 12 or 11 times higher for children whose Oppositional or Withdrawn scale scores were above the 95th percentile point than for children whose scores were in the normal range, respectively. Chi-square tests indicated that significantly more children from the referred than from the nonreferred sample scored above the cutpoint on all scales ($p < .01$). On basis of having a Total Problems score above the 85th percentile cutpoint, 73.3% of the children could be correctly classified as being referred.

Predictive Correlations with CBCL/4-18 Scores at Ages 4 to 5

The stability of parent-reported behavior problems as assessed with the CBCL/2-3 was assessed over a 2-year period by computing Pearson correlations between problem scores obtained for the community sample for the CBCL/2-3 with the corresponding CBCL/4-18 scores obtained at follow-up (Table IV). On average, the correlations between Time 1 and Time 2 Internalizing, Externalizing, and Total Problems scores were .41, .54, and 60, respectively. Both Externalizing and Total Problems scores showed significantly higher stabilities than Internalizing scores ($p < .05$, according to Fisher's z-test). Although 3-year-olds and girls tended to have somewhat more stable scores, none of the sex- and age-related differences were significant.

Convergent and Divergent Validity

An indication of the construct validity of the CBCL/2-3 would be if the ratings of problem behavior showed stronger relations with a measure which taps the same aspects of functioning, i.e., problem behavior, than with a measure which taps other aspects of functioning, i.e., problem behavior, than with a measure which taps other aspects of function, such as general developmental level. To test this, we computed the concurrent Pearson correlations of the CBCL/2-3 scores with the BCL score and with the MCDI General Development score. Like the CBCL/2-3, both the BCL and the MCDI are parent rating instruments, i.e., all three measures used the same

Table V. Number of Items in Common and Pearson Correlations Between Dutch and American Versions of CBCL/2-3 Scales[a]

Dutch scales	American scales	Number of items in common	Correlation
Oppositional	Aggressive Behavior	9	.94
Withdrawn/Depressed	Withdrawn	6	.88
Aggressive	Destructive Behavior	4	.82
	Aggressive Behavior	4	.80
Anxious	Anxious/Depressed	5	.84
Overactive	—	—	—
Sleep problems	Sleep problems	7	1.00
Internalizing	Internalizing	14	.90
Externalizing	Externalizing	20	.97

[a] n = 420. CBCL/2-3 = Child Behavior Checklist for Ages 2-3. American scales are derived from Achenbach (1992). All correlations were significant at $p < .001$.

method of data collection. Therefore, relations in the expected direction of CBCL/2-3 scores with BCL and MCDI scores, respectively, may be regarded as evidence of construct validity for the CBCL/2-3 (Campbell & Fiske, 1959). We expected that the correlations of CBCL/2-3 scores with BCL scores, measuring levels of problem behavior, would be higher than the correlations with MCDI scores, measuring developmental levels. Since the BCL was developed for children aged 3 years, the relation between CBCL/2-3 and BCL was only analyzed for children in this age group (n = 207).

All correlations between the BCL score and the CBCL/2-3 scale scores and Total Problems score (see Table II) were significant at $p < .01$, with a mean r of .41, ranging from $r = .65$ for Total Problems to $r = .28$ for Withdrawn/Depressed. The mean rs were not different for boys and girls.

The BCL was originally designed as a screening instrument to identify preschoolers at risk for behavior problems. Of the 3-year-olds studied by Richman (1977), 78.5% could be correctly classified using a BCL score ≥ 10 and a clinical rating of mild, moderate, or severe behavior problems as criterion. The CBCL/2-3 might be applied for the same purpose using a clinical cutoff point, e.g., a score above the 85th percentile on Total Problems as proposed by Achenbach (1992). We compared the agreement between scores > 85th percentile on the CBCL/2-3 Total Problems score and BCL scores ≥ 10. Of the children with BCL scores ≥ 10, 28.6% had a CBCL/2-3 scores below the 85th percentile point, and 10.7% of the children with low BCL scores had high CBCL/2-3 scores, producing a 12.4% overall misclassification rate. Thus, having a high BCL score could be predicted from the CBCL/2-3 score for more than 87% of the cases. The relative risk odds ratio (Fleiss, 1981) for having a high BCL score relative to having a high CBCL/2-3 score was 21.0, which means that children who had high CBCL/2-3 scores had 21 times greater odds of having high BCL scores than children who had low CBCL/2-3 scores.

Five out of nine correlations of CBCL/2-3 scores with the MCDI General Development score were significant ($p < .05$). However, correlations were very low, with absolute values ranging from .03 to .16 (mean $r = -.06$; see Table II), indicating minimal relationships between CBCL/2-3 problem scores and the General Development score.

Comparison with American Syndromes

Because comparability of methods for the assessment of psychopathology across studies is important for integrating findings from different studies into a cumulative knowledge base, and for detecting cross-cultural differences and similarities in psychopathology (cf. Achenbach, Verhulst, et al., 1987), we compared the syndromes obtained in the present study with those reported by Achenbach (1992).

As a quantitative test of the similarity of the syndromes across cultures, for each child in the matched referred and nonreferred sample, scores were computed by summing the raw scores of 0, 1, and 2 on each CBCL/2-3 item belonging to each American syndrome scale. Pearson correlations were computed between the children's raw scores on the Dutch and American syndrome scales. Table V gives the correlation coefficients between scores on the Dutch and American scales that were the most similar, as well as

the number of corresponding items in these scales. The correlations between similar scales were all high and significant ($p < .001$), and all were significantly higher than those between any other combinations of Dutch and American syndrome scales (except Internalizing and Externalizing), with the smallest $z = 2.50$ ($p < .01$) for the difference between the highest r and the second highest r for each syndrome.

DISCUSSION

The aim of this study was to assess the applicability of the Dutch version of the Child Behavior Checklist for Ages 2-3, since until now no validated measures of problem behavior existed in the Netherlands. A series of exploratory and confirmatory factor analyses on CBCL/2-3 scores for three different samples indicated a seven-factor model for all three samples. The syndrome scales derived from these factors were labeled Oppositional, Withdrawn/Depressed, Aggressive, Anxious, Overactive, Sleep Problems, and Somatic Problems. Factor intercorrelations and a second-order factor analysis provided support for two broad-band groupings of problem behaviors — Externalizing and Internalizing.

The internal consistency and test-retest stability coefficients were quite acceptable for most of the CBCL/2-3 scales, and comparable to Achenbach's (1992) findings. Interparent agreement for the scale scores was generally moderate, but quite comparable to the levels of agreement found for the American version of the CBCL/2-3 (Achenbach, 1992), as well as for other cross-informant comparisons (Achenbach, McConaughy, & Howell, 1987). The interparent agreement for scales reflecting behavior in the Internalizing spectrum was lower than for Externalizing scales, which compared to Achenbach's (1992) findings for the 3-year-old age group.

The factor structure of the CBCL/2-3 resulting from this study was fairly invariant across different Dutch samples. Correlations of scale scores using the Dutch and American syndromes showed high concordance at the level of scores across cultures for all scales retained from the analyses of the Dutch version. Both the high congruency of the factor structure across different samples and the similarity of most scales across cultures are supportive of the validity of the differentiation of young children's problem behaviors into syndromes. These syndromes reflect oppositional behavior, aggression, and overactive behavior, constituting an externalizing grouping, and anxiety and withdrawal, constituting an internalizing groping. Sleep problems represent a separate syndrome at this age. Although we obtained a somatic problems factor comparable to the American one from our exploratory analyses, it appeared to be not robust enough to retain as a scale. This may partly be due to the relatively low frequencies of reported somatic problems without medical cause in our samples.

Unexpectedly, we found a distinct overactive syndrome. Until now, a separate syndrome including items indicative of hyperactivity and attention problems in young preschoolers had only been found in one other study using rating scales. Behar and Stringfield (1974) obtained a distinct Hyperactive-Distractable factor using a teacher-rated questionnaire. However, only one-sixth of their sample was 3 years old. The present study thus confirmed the possibility of a further differentiation of toddler's externalizing problem behaviors using parent ratings.

The convergence between American and Dutch syndromes emerged despite differences in the factor-analytic approach, as might be expected with large numbers of variables (Snook & Gorsuch, 1989). The use of factor analysis instead of principal-components analysis made it possible to employ techniques that provided better structure and more reliable estimates of factor loadings, which were of value in comparisons across samples. This approach yielded highly discriminating scales, which also resulted in a lower correlation (.40) between the Internalizing and the Externalizing scales than was obtained for similar scales of the American version of the CBCL/2-3 (.75), and for the American and Dutch versions of the CBCL/4-18. However, in using a different factor-analytic technique as well as differently composed samples than those used by Achenbach (1992), part of the cross-cultural comparability of results was lost. In other words, we are not able to decide to what extent the somewhat different factor structure of the CBCL/2-3 obtained in this study compared to the one found for American samples was the result of the use of different factor-analytic methods and different samples, or was a reflection of true cross-cultural differences in young preschoolers' problem behaviors. However, these differences, even if they reflect cultural variation, appear to be relatively trivial, given the large degree of overlap among the symptoms of the various syndromes.

The construct validity of the CBCL/2-3 was confirmed by its significant associations with the Richman Behavior Checklist using both continuous and categorical scores. We found a correlation between the BCL and the Total Problems score of .65, and a concordance between categorical scores of 87%. Further, all correlations between the MCDI General Development and CBCL/2-3 scores were low (with a mean of –.06). The evidence of construct validity of the Dutch CBCL/2-3 is limited by the absence of other validated measures of behavior problems for young preschoolers that use methods other than parental ratings.

The criterion-related validity of the CBCL/2-3 was supported by the significant differences between demographically matched referred and nonreferred children's scale scores, which were highly comparable to the American findings reported by Achenbach (1992). Only a moderate effect of referral status was found for the Anxious scale, however. The total correct classification rate of 73.3% using the Total Problems score was somewhat lower than the 80.1% found for CBCL Total Problems scores for a sample of Dutch children aged 4 to 16 (Verhulst et al., 1985).

Previous American-Dutch cross-cultural comparisons of CBCL syndromes indicated a high degree of congruence between the American and Dutch scale structures for boys and girls aged 4 to 18 years (Achenbach, Verhulst, et al., 1987; de Groot, Koot, & Verhulst, 1994; Verhulst, et al., 1988). Our current results for the age group of 2- to 3-year-olds correspond with these earlier findings, demonstrating only small cross-cultural differences between American and Dutch scale structures, despite considerable methodological differences in the derivation of problem scales.

In conclusion, our findings indicated that the Dutch CBCL/2-3 shows psychometric properties comparable to those obtained for American samples. This supports the cross-cultural validity of the CBCL/2-3 and its use in clinical practice and for future research.

ACKNOWLEDGMENTS

This research was supported by the Dutch National Programme for Stimulation of Health Research (SGO), the Dutch Ministry of Well-Being, Health, and Cultural Affairs (WVC), and the Rotterdam Foundation for Autism and Congenital Contact Disorders (SAR).

REFERENCES

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile.* Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist/2-3 and 1992 profile.* Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1993). *Empirically based taxonomy: How to use syndromes and profile types derived from the CBCL, TRF, and YSR.* Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., Edelbrock, C., & Howell, C. T. (1987). Empirically based assessment of the behavioral/emotional problems of 2- and 3-year-old children. *Journal of Abnormal Child Psychology, 15,* 629-650.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213-232.

Achenbach, T. M., Verhulst, F. C., Baron, G. D., & Althaus, M. (1987). A comparison of syndromes derived from the Child Behavior Checklist for American and Dutch boys aged 6-11 and 12-16. *Journal of Child Psychology and Psychiatry, 28,* 437-453.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Behar, L. B., & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology, 10,* 601-610.

Campbell, S. B. (1990). *Behavior problems in preschool children: Clinical and developmental issues.* New York: Guilford Press.

Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry, 36,* 113-149.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Central Bureau of Statistics (CBS). (1987). *Standard educational classification 1978; Edition 1987.* Voorburg: CBS.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

de Groot, A., Koot, H. M., & Verhulst, F. C. (1994). Cross-cultural generalizability of the Child Behavior Checklist Cross-informant syndromes. *Psychological Assessment, 6,* 225-230.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

Ireton, H., & Thwing, E. J. (1974). *Minnesota Child Development Inventory.* Minneapolis: Behavior Science Systems.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications.* Chicago: SPSS.

Koot, H. M., & Verhulst, F. C. (1991). Prevalence of problem behavior in Dutch children aged 2-3. *Acta Psychiatrica Scandinavica, 83* (Suppl. 367), 1-37.

Richman, N. (1977). Is a behaviour checklist for preschool children useful? In P. J. Graham (Ed.), *Epidemiological approaches to child psychiatry* (pp. 125-137). London: Academic Press.

Richman, N., & Graham, P. J. (1971). A behavioural screening questionnaire for use with three-year-old children: Preliminary findings. *Journal of Child Psychology and Psychiatry, 12,* 5-33.

Richman, N., & Lansdown, R. (Eds.). (1988). *Problems of preschool children.* Chichester: Wiley.

Richman, N., Stevenson, J., & Graham, P. J. (1982). *Pre-school to school: A behavioural study.* London: Academic Press.

Rutter, M., Tuma, A. H., & Lann, I. S. (Eds.). (1988). *Assessment and diagnosis in child psychopathology.* New York: Guilford Press.

SAS Institute (1989). *SAS user's guide, Release 6.07.02.* Cary, NC: Author.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin, 106,* 148-154.

Spiker, D., Kraemer, H. C., Constantine, N. A., and Bryant, D. (1992). Reliability and validity of behavior problem checklists as measures of stable traits in low birth weight, premature preschoolers. *Child Development, 63,* 1481-1496.

Tanaka, J. S., & Huba, G. J. (1984). Hierarchical confirmatory factor analyses of psychological distress measures. *Journal of Personality and Social Psychology, 46,* 621-635.

Trad, P. V. (1989). *The preschool child.* New York: Wiley.

Van Westerlaak, J. M., Kropman, J. A., & Collaris, J. W. M. (1975). *Beroepenklapper* [Index of occupations]. Nijmegen, The Netherlands: Instuut voor Toegepaste Sociologie.

Verhulst, F. C. (1995). A review of community studies. In F. C. Verhust & H. M. Koot (Eds.), *The epidemiology of child and adolescent psychopathology* (pp. 146-177). Oxford, England: Oxford University Press.

Verhulst, F. C., Achenbach, T. M., Althaus, M., & Akkerhuis, G. W. (1988). A comparison of syndromes derived from the Child Behavior Checklist for American and Dutch girls aged 6-11 and 12-16. *Journal of Child Psychology and Psychiatry, 29,* 879-895.

Verhulst, F. C., Akkerhuis, G. W., & Althaus, M. (1985). Mental health in dutch children: I. A cross-cultural comparison. *Acta Psychiatrica Scandinavica, 72* (Suppl. No. 323, pp. 1-108).

Wenar, C. (1994). *Developmental psychopathology: From infancy through adolescence.* New York: McGraw-Hill.